What is claimed is:

1.   A method of pattern recognition, comprising:

training a plurality of observable operator models for a plurality of

5   characteristic events, the observable operator models comprising a plurality of

observable operators;

receiving an unknown input;

computing a plurality of matching probabilities, one matching probability

for each one of the plurality of characteristic events using the plurality of

10   observable operators, wherein each matching probability is a probability that the

unknown input matches a particular characteristic event;

selecting a maximum matching probability from the plurality of matching

probabilities; and

displaying a characteristic event having the maximum matching

15   probability.

2.   The method as recited in claim 1, wherein the unknown input occurs at a

particular point in an input stream, the input stream comprising a sequence

occurring prior to the unknown input.

20

3.   The method as recited in claim 2, each matching probability is a

probability that the unknown input matches a particular characteristic event,

given the sequence occurring prior to the unknown input.

25   4.   The method as recited in claim 1, wherein the unknown input is a word

and the characteristic events define a vocabulary.

5.   The method as recited in claim 1, wherein the training act comprises:

computing a probability transition matrix;

30   computing an occurrence count matrix;

estimating the plurality of observable operators from the probability

transition matrix and the occurrence count matrix; and

standardizing the plurality of observable operators.

6. The method as recited in claim 5, wherein standardizing the plurality of observable operators further comprises computing a mean and standard deviation for each observable operator.

7. A computer-readable medium having computer-executable instructions for performing a method of recognizing speech, the method comprising:

sampling a first input stream, resulting in a plurality of samples;

performing a spectral analysis of the samples to obtain a plurality of feature vectors;

clustering the feature vectors to form a plurality of observation vectors;

training at least one observable operator model using the observation vectors; and

recognizing at least one part of speech from a second input stream using the at least one observable operator model.

8. The computer-readable medium as recited in claim 7, wherein the spectral analysis comprises:

removing noise from the samples;

performing pre-emphasis on the samples in order to spectrally flatten the samples;

blocking the spectrally flattened samples into framed samples;

windowing the framed samples to obtain windowed samples with signal energy substantially concentrated at the center of the frames;

performing auto-correlation analysis for each windowed sample to obtain auto-correlated samples; and

performing linear predictive coding (LPC) analysis for each auto-correlated sample to obtain feature vectors.

9. The computer-readable medium as recited in claim 7, wherein the at least one observable operator model is stateless.

10.   The computer-readable medium as recited in claim 9, wherein a probability of selecting an operator is computed using the operator itself.

11.   The computer-readable medium as recited in claim 10, wherein the computational efficiency is about $O(n + k)$ where n is a number of samples and k is a constant.

12.   The computer-readable medium as recited in claim 10, wherein the at least one part of speech comprises a concatenation of a plurality of operators.

13.   A data structure of an observable operator model used to recognize patterns, comprising:

a plurality of characteristic events corresponding to a input stream, the input stream comprising a plurality of stream elements and the input stream comprising a plurality of sequences;

an initial distribution vector, wherein each element of the initial distribution vector comprises a particular probability that a particular characteristic event is an initial event, the particular characteristic event being one of the plurality of characteristic events;

a probability transition matrix, wherein each element of the probability transition matrix comprises an estimate of a particular probability of producing a particular characteristic event, after observing a particular sequence;

an occurrence count matrix, wherein each element of the occurrence count matrix comprises an estimate of a particular probability of producing the particular characteristic event, after observing a particular stream element followed by the particular sequence; and

at least one observable operator calculable from the probability transition matrix and the occurrence count matrix;

wherein the plurality of characteristic events, the initial distribution vector, the probability transition matrix, the occurrence count matrix, and the at

least one observable operator are storable on a storage medium during a training phase and retrievable during a recognition phase.

14.   The data structure as recited in claim 13, wherein each element of the occurrence count matrix comprises a calculation, during the training phase, of how frequently the particular characteristic event occurs after observing the particular stream element followed by the particular sequence.

15.   The data structure as recited in claim 13, wherein each element of the occurrence count matrix comprises a number of occurrences of the particular stream element followed by the particular sequence, the number being countable during the training phase.

16.   The data structure as recited in claim 13, wherein each element of the probability transition matrix comprises a calculation, during the training phase, of how frequently the particular characteristic event occurs, after observing the particular sequence.

17.   The data structure as recited in claim 13, wherein each element of the probability transition matrix comprises a number of occurrences of the particular sequence, the number being countable during the training phase.

18.   The data structure as recited in claim 13, wherein the observable operator (τ) is a linear, observable operator.

19.   The data structure as recited in claim 13, wherein the columns of the probability transition matrix sum to 1.

20.   The data structure as recited in claim 19, wherein the elements of the initial distribution vector sum to 1.

21. The data structure as recited in claim 20, wherein at least one element of the matrices and vectors is a negative value.

22. The data structure as recited in claim 13, wherein the at least one observable operator is equal to the inverse of the probability transition matrix times an element of the occurrence count matrix.

23. A computer-readable medium having computer-executable instructions for performing a method for modeling a process with the data structure of the observable operator model as recited claim 13, the method comprising:

creating at least one data structure of the observable operator model;

storing the plurality of characteristic events of the at least one data structure of the observable operator model;

storing the initial distribution vector of the at least one data structure of the observable operator model;

storing the probability transition matrix of the at least one data structure of the observable operator model;

storing the occurrence count matrix of the at least one data structure of the observable operator model; and

storing the at least one observable operator of the at least one data structure of the observable operator model.

24. The computer-readable medium as recited in claim 23, the method further comprising:

reading the plurality of characteristic events of the at least one data structure of the observable operator model;

reading the initial distribution vector of the at least one data structure of the observable operator model;

reading the probability transition matrix of the at least one data structure of the observable operator model;

reading the occurrence count matrix of the at least one data structure of the observable operator model; and

reading the at least one observable operator of the at least one data
structure of the observable operator model.

25. A method for recognizing speech, comprising:

sampling an input signal;

converting the input signal to a discrete signal;

storing the discrete signal in a buffer;

reading a frame of data from the buffer;

checking for silence or noise in the frame;

removing any silence and noise from the frame;

spectrally flattening a signal in the frame;

performing frame windowing on the frame;

computing a moving weighted average for the frame;

performing feature extraction on the frame using a mathematical model;

clustering the frame with previously read frames;

training a plurality of observable operator models;

recognizing at least one unknown word using the observable operator
models; and

displaying a recognized word corresponding to the at least one unknown
word.

26. The method as recited in claim 25, wherein reading the frame of data
from the buffer is performed repeatedly, until the input signal is exhausted.

27. The method as recited in claim 25, further comprising:

discarding the frame and reading a next frame; and

introducing overlap for a next read.

28. The method as recited in claim 25, wherein training the plurality of
observable operator models comprises:

computing a transition between observables in the observation sequence;
and

computing an estimate of the observables.

29.   The method as recited in claim 28, wherein computing the transition between observables in the observation sequence is performed by computing a

5   probability of observing possible subsequences among the observables.

30.   The method as recited in claim 29,wherein computing an estimate of the observables comprises:
        developing the plurality of observable operator models; and

10   structuring the probability to produce substantially well-defined linear operators for the plurality of observable operator models.

31.   The method as recited in claim 30, further comprising:
        standardizing the plurality of observable operator models such that

15   variation between similar signals is substantially minimized; and
        storing the plurality of observable operator models in a system repository.

32.   The method as recited in claim 30, wherein recognizing the at least one

20   unknown word using the plurality of observable operator models comprises:
        determining a distribution of the observables;
        computing a probability for each of the plurality of observable operator models in the system repository; and
        selecting a most probable one of the plurality of observable operator

25   models as the at least one recognized word.

33.   The method as recited in claim 25, wherein spectrally flattening the signal is performed using a first order filter.

30   34.   The method as recited in claim 25, wherein performing frame windowing is performed by multiplying the signal in the frame by a window so that

information in the signal is concentrated substantially towards a center of the frame.

35. The method as recited in claim 34, wherein the window is selected from the group consisting of a Hamming window and a Hanning window.

36. The method as recited in claim 25, wherein computing the moving weighted average for the frame is performed using auto-correlation.

37. The method as recited in claim 25, wherein the mathematical model is linear predictive coding.

38. The method as recited in claim 37, wherein the linear predictive coding models a vocal tract.

39. The method as recited in claim 25, wherein clustering the frame with previously read frames comprises grouping similar features.

40. The method as recited in claim 39, wherein clustering the frame with previously read frames comprises obtaining an observation sequence.

41. The method as recited in claim 40, wherein obtaining the observation sequence comprises obtaining indices of the clusters.